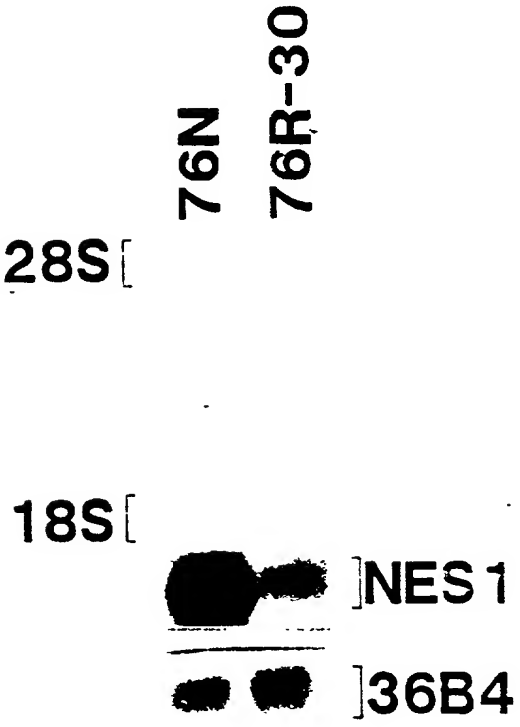

FIG. 1

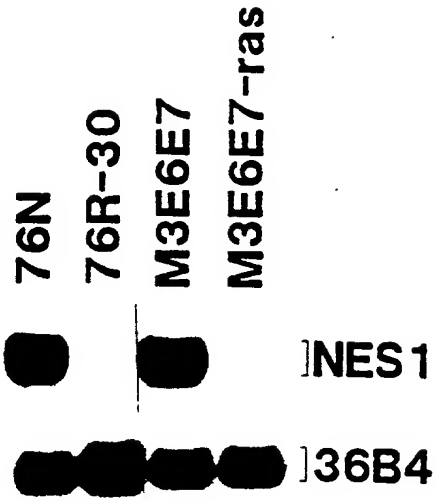

FIG. 3

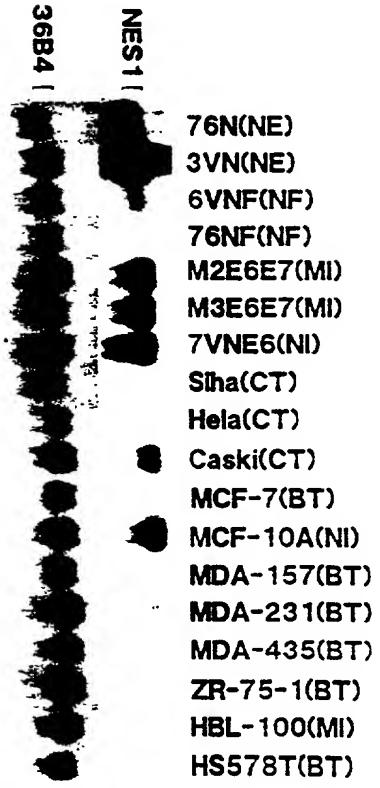

FIG. 2A

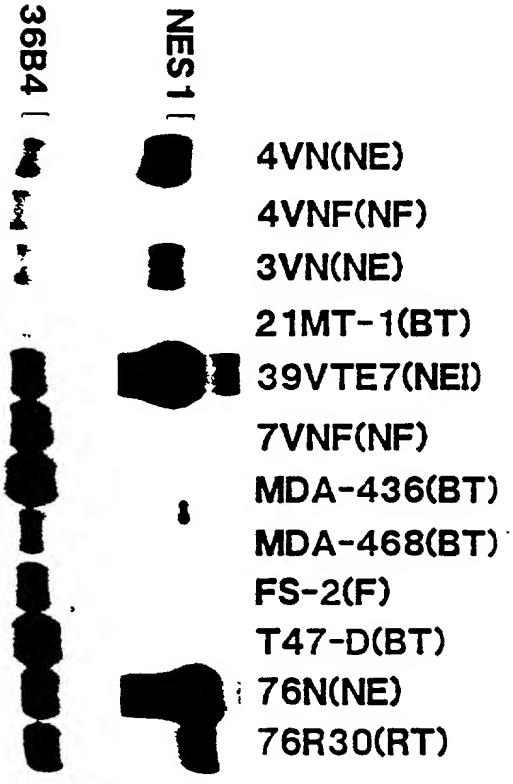

FIG. 2B

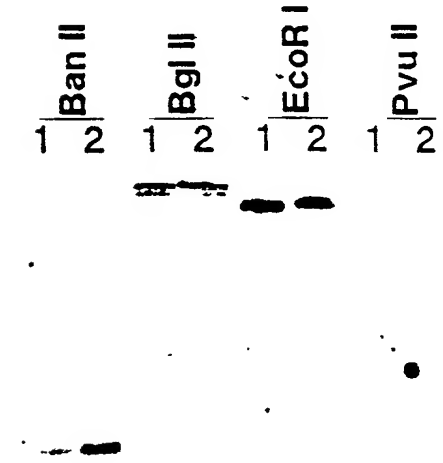

FIG. 4

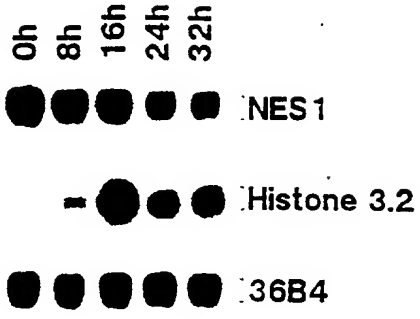

FIG. 5A

FIG. 6

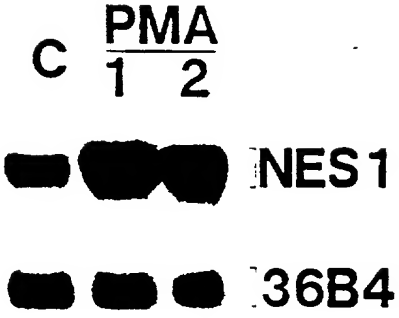

FIG. 8

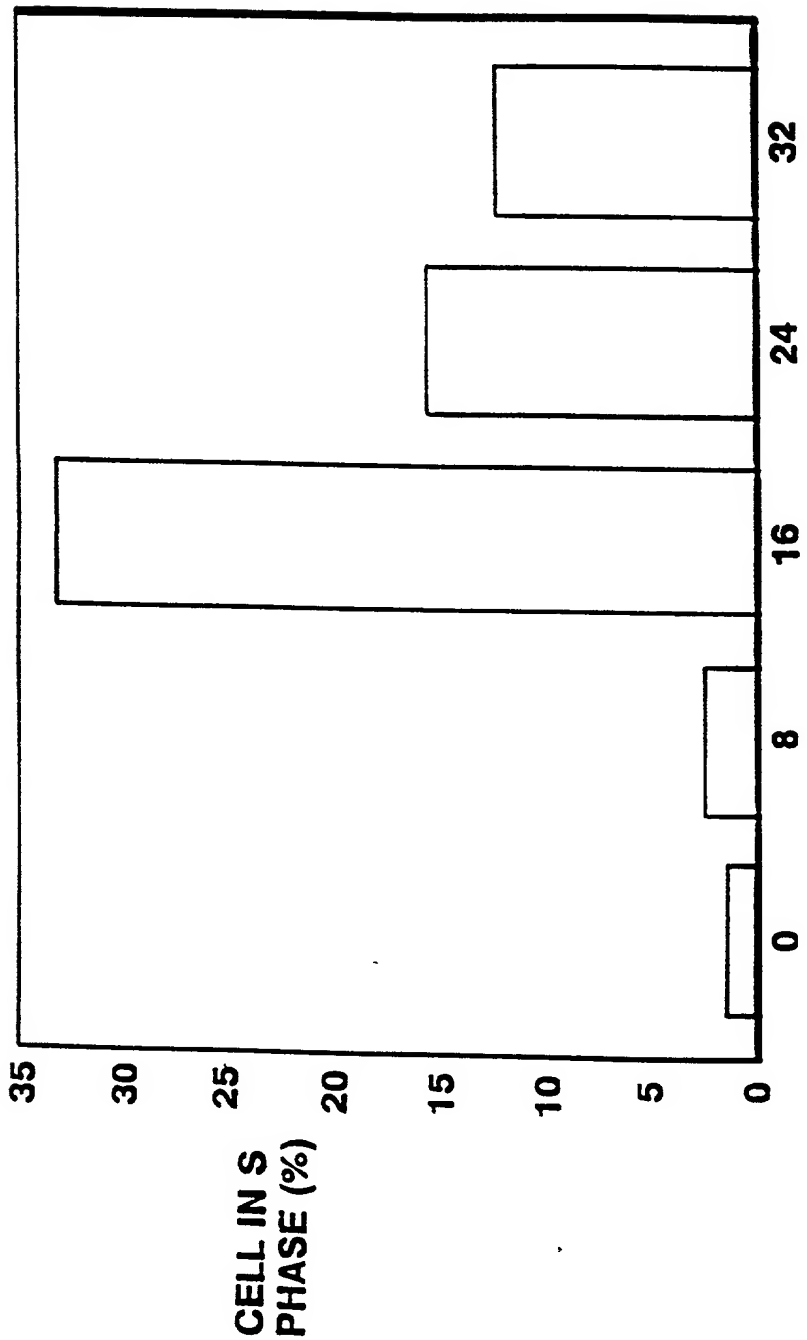

FIG. 5B

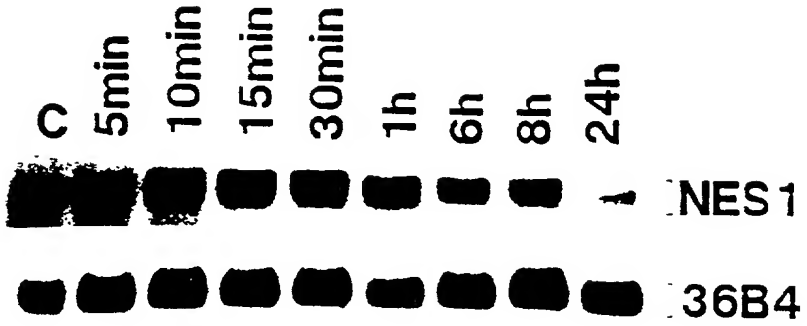
C
5min
10min
15min
30min
1h
6h
8h
24h
NES1
36B4

FIG. 7A

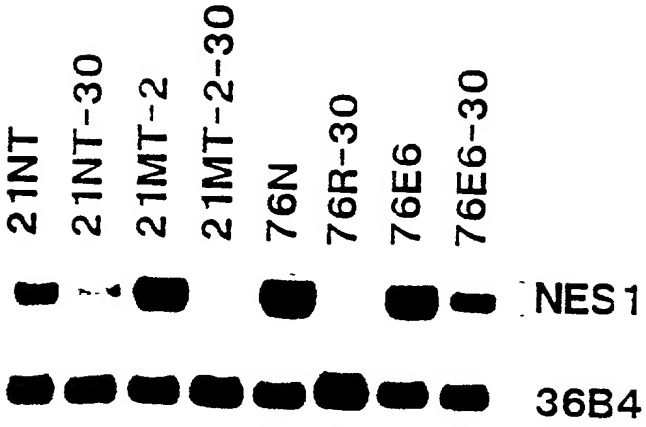
21NT
21NT-30
21MT-2
21MT-2-30
76N
76R-30
76E6
76E6-30
NES1
36B4

FIG. 7B

Comparison of Sequences near putative active site

	86	137	223	224	227	228	229	230	231	241	243	245
NES1	H	D	D	P	S	D	S	G	G	G	L	W
Human Pancreatic Trypsinogen III	H	D	D	S	R	D	S	G	G	G	V	W
	▲	▲	△	△			▲			△	△	△

▲: Catalytic triad

△: Residues important for substrate binding and specificity

FIG. 9A

Alignment of NES1 amino acid sequence with other serine proteases

```
           1                                                  50
  Mmtryar  ..MSALLILA LVGA...... .......... AVAFPVDD.D DKIVG...GY
 Hstryivb  .....LELHP LLGGRTWRAA RDADGCEALG TVAVPFDD.D DKIVG...GY
 Rntrypvb  ..MKICIFFT LLGT...... .......... VAAFPTEDND DRIVG...GY
 Sstrypii  .......... .......... .......... .AAFATE..D DKIVG...GY
     Nes1  MRAPHLHLSA ASGARALAKL LPLLMAQLWA AEAALLPQND TRLDPEAYGA

           51                                                100
  Mmtryar  TCRESSVPYQ VSLNAGYHF. CGGSLINDQW VVSAAHCYKY RIQVRLGEHN
 Hstryivb  TC.ENSLPYQ VSLNSGSHF. CGGSLISEQW VVSAAHCYKT RIQVRLGEHN
 Rntrypvb  TCQEHSVPYQ VSLNAGSHI. CGGSLITDQW VLSAAHCYHP QLQVRLGEHN
 Sstrypii  ECKAYSQPHQ VSLNSGYHF. CGGSLVNENW VVSAAHCYQS RVEVRLGEHN
     Nes1  PCARGSQPWQ VSLFNGLSFH CAGVLVDQSW VLTAAHCGNK PLWARVGDDH

           101                                               150
  Mmtryar  INVLEGNEQF VDSAKIIRHP NYN....... .SWTLDNDIM LIKLASPVTL
 Hstryivb  IKVLEGNEQF INAAKIIRHP KYN....... .RDTLDNDIM LIKLSSPAVI
 Rntrypvb  IYEIEGAEQF IDAAKMILHP DYD....... .KWTVDNDIM LIKLKSPATL
 Sstrypii  IQVTEGSEQF ISSSRVIRHP NYS....... .SYNIDNDIM LIKLSKPATL
     Nes1  LLLLQG.EQL RRTTRSVVHP KYHQGSGPIL PRRTDEHDLM LLKLARPVVP
                                 EXTRA INSERT
           151                                               200
  Mmtryar  NARVASVPLP SSCAPAGTQC LISGWGNTLS NGVNNPDLLQ CVDAPVLPQA
 Hstryivb  NARVSTISLP TAPPAAGTEC LISGWGNTLS FGADYPDELK CLDAPVLTQA
 Rntrypvb  NSKVSTIPLP QYCPTAGTEC LVSGWG.VLK FGFESPSVLQ CLDAPVLSDS
 Sstrypii  NTYVQPVALP TSCAPAGTMC TVSGWGNTMS STAD.KNKLQ CLNIPILSYS
     Nes1  GPRVRALQLP YRCAQPGDQC QVAGWGTTAA RRVKYNKGLT CSSITILSPK

           201                                               250
  Mmtryar  DCEASYPGDI TNNMICVGFL EGGKDSCQGD SGGPVVCNGE LQGIVSWG.Y
 Hstryivb  ECKASYPGKI TNSMFCVGFL EGGKDSCQRD SGGPVVCNGQ LQGVVSWG.H
 Rntrypvb  VCHKAYPRQI TNNMFCLGFL EGGKDSCQYD SGGPVVCNGE VQGIVSWG.D
 Sstrypii  DCNNSYPGMI TNAMFCAGYL EGGKDSCQGD SGGPVVCNGE LQGVVSWG.Y
     Nes1  ECEVFYPGVV TNNMICAG.L DRGQDPCQSD SGGPLVCDET LQGILSWGVY

           251                           280
  Mmtryar  GCAQPDAPGV YTKVCNYVDW IQNTIADN*.
 Hstryivb  GCAWKNRPGV YTKVYNYVDW IKDTIAANS*
 Rntrypvb  GCALEGKPGV YTKVCNYLNW IQQTVAAN*.
 Sstrypii  GCAEPGNPGV YAKVCIFNDW LTSTMATY*.
     Nes1  PCGSAQHPAV YTQICKYMSW INKVIRSN*.
```

FIG. 9B

```
  1  MRAPHLHLSA ASGARALAKL LPLLMAQLWA AEAALLPQND TRLDPEAYGA
 51  PCARGSQPWQ VSLFNGLSFH CAGVLVDQSW VLTAAHCGNK PLWARVGDDH
101  LLLLQGEQLR RTTRSVVHPK YHQGSGPILP RRTDEHDLML LKLARPVVPG
151  PRVRALQLPY RCAQPGDQCQ VAGWGTTAAR RVKYNKGLTC SSITILSPKE
201  CEVFYPGVVT NNMICAGLDR GQDPCQSDSG GPLVCDETLQ GILSWGVYPC
251  GSAQHPAVYT QICKYMSWIN KVIRSN* (SEQ ID NO: 1)
```

FIG. 10

```
   1  ACCAGCGGCA GACCACAGGC AGGGCAGAGG CACGTCTGGG TCCCCTCCCT
  51  CCTTCCTATC GGCGACTCCC AGATCCTGGC CATGAGAGCT CCGCACCTCC
 101  ACCTCTCCGC CGCCTCTGGC GCCCGGGCTC TGGCGAAGCT GCTGCCGCTG
 151  CTGATGGCGC AACTCTGGGC CGCAGAGGCG GCGCTGCTCC CCCAAAACGA
 201  CACGCGCTTG GACCCCGAAG CCTATGGCGC CCCGTGCGCG CGCGGCTCGC
 251  AGCCCTGGCA GGTCTCGCTC TTCAACGGCC TCTCGTTCCA CTGCGCGGGT
 301  GTCCTGGTGG ACCAGAGTTG GGTGCTGACG GCCGCGCACT GCGGAAACAA
 351  GCCACTGTGG GCTCGAGTAG GGGATGATCA CCTGCTGCTT CTTCAGGGCG
 401  AGCAGCTCCG CCGGACGACT CGCTCTGTTG TCCATCCCAA GTACCACCAG
 451  GGCTCAGGCC CCATCCTGCC AAGGCGAACG GATGAGCACG ATCTCATGTT
 501  GCTAAAGCTG GCCAGGCCCG TAGTGCCGGG GCCCCGCGTC CGGGCCCTGC
 551  AGCTTCCCTA CCGCTGTGCT CAGCCCGGAG ACCAGTGCCA GGTTGCTGGC
 601  TGGGGCACCA CGGCCGCCCG GAGAGTGAAG TACAACAAGG GCCTGACCTG
 651  CTCCAGCATC ACTATCCTGA GCCCTAAAGA GTGTGAGGTC TTCTACCCTG
 701  GCGTGGTCAC CAACAACATG ATATGTGCTG GACTGGACCG GGGCCAGGAC
 751  CCTTGCCAGA GTGACTCTGG AGGCCCCCTG GTCTGTGACG AGACCCTCCA
 801  AGGCATCCTC TCGTGGGGTG TTTACCCCTG TGGCTCTGCC CAGCATCCAG
 851  CTGTCTACAC CCAGATCTGC AAATACATGT CCTGGATCAA TAAAGTCATA
 901  CGCTCCAACT GATCCAGATG CTACGCTCCA GCTGATCCAG ATGTTATGCT
 951  CCTGCTGATC CAGATGCCCA GAGGCTCCAT CGTCCATCCT CTTCCTCCCC
1001  AGTCGGCTGA ACTCTCCCCT TGTCTGCACT GTTCAAACCT CTGCCGCCCT
1051  CCACACCTCT AAACATCTCC CCTCTCACCT CATTCCCCCA CCTATCCCCA
1101  TTCTCTGCCT GTACTGAAGC TGAAATGCAG GAAGTGGTGG CAAAGGTTTA
1151  TTCCAGAGAA GCCAGGAAGC CGGTCATCAC CCAGCCTCTG AGAGCAGTTA
1201  CTGGGGTCAC CCAACCTGAC TTCCTCTGCC ACTCCCCGCT GTGTGACTTT
1251  GGGCAAGCCA AGTGCCCTCT CTGAACCTCA GTTTCCTCAT CTGCAAAATG
1301  GGAACAATGA CGTGCCTACC TCTTAGACAT GTTGTGAGGA GACTATGATA
1351  TAACATGTGT ATGTAAATCT TCATGTGATT GTCATGTAAG GCTTAACACA
1401  GTGGGTGGTG AGTTCTGACT AAAGGTTACC TGTTGTCGTG AAAAAAAAAA
1451  AAAA  (SEQ ID NO: 2)
```

FIG. 11